

Chapter 7 - Univariate and Descriptive Statistics

Supplemental Code

Brian Fogarty

This document provides the code for creating the SAT, Wayne County, and St. Louis City histograms that appear in Chapter 7.

SAT Histogram

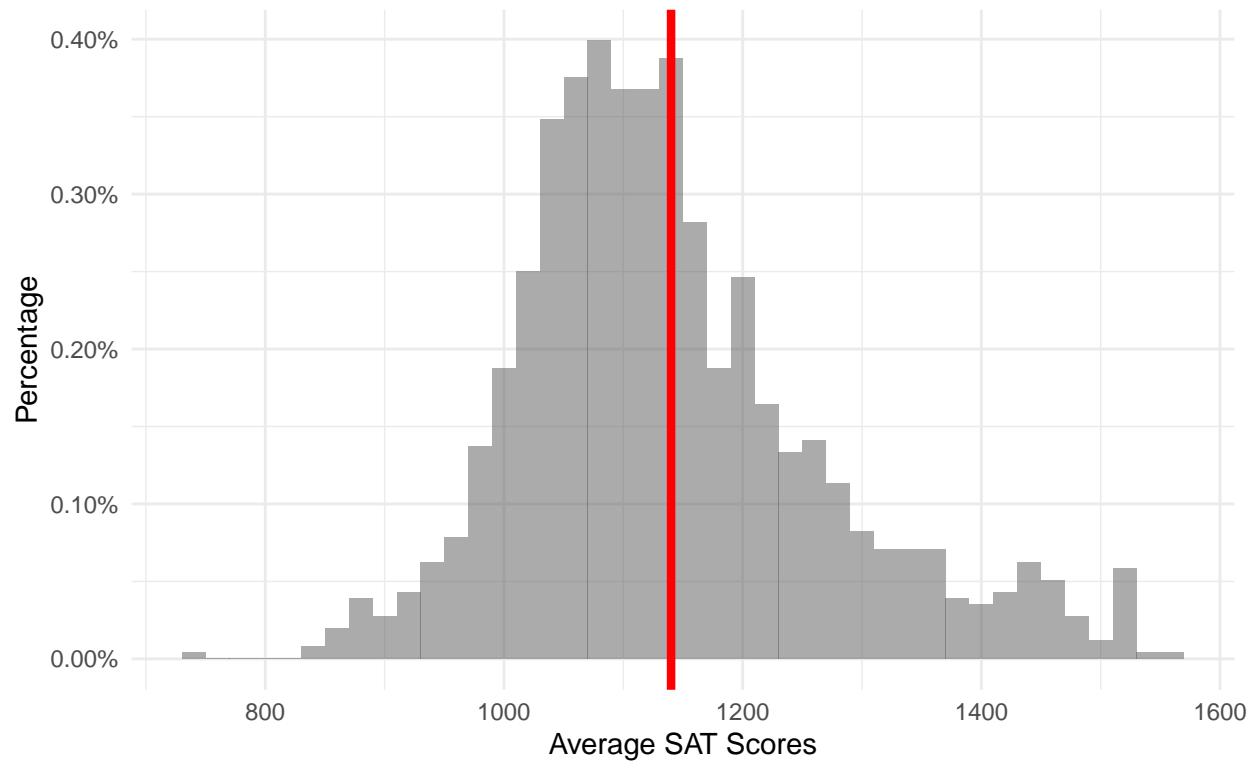
```
library(tidyverse)
library(scales)

sat <- read_csv("sat_unis.csv")
glimpse(sat)

Rows: 1,278
Columns: 2
$ uni      <chr> "Alabama A & M University", "University of Alabama at Birmingh~
$ sat_avg  <dbl> 939, 1234, 1319, 946, 1261, 1082, 1300, 1230, 1066, 1076, 1084~

sat %>%
ggplot() +
  geom_histogram(mapping = aes(sat_avg, y = stat(density)),
                 binwidth = 20, alpha = 1/2) +
  theme_minimal() +
  scale_y_continuous(labels = percent_format()) +
  geom_vline(xintercept = 1140, size = 1.5, colour = "red") +
  labs(x="Average SAT Scores", title="Average SAT Scores Among U.S. Universities",
       y="Percentage", caption="Red vertical line is the mean.")
```

Average SAT Scores Among U.S. Universities



Red vertical line is the mean.

Wayne County, MI Histogram

```
library(readxl)
wayne <- read_xlsx("wayne_voting_precincts.xlsx")
glimpse(wayne)
```

```
Rows: 981
Columns: 3
$ Label      <chr> "Voting District 1632200005125, Wayne County, Michigan", "Vo~
$ white_pct  <dbl> 8.450704, 62.903226, 76.329114, 75.844516, 76.531165, 83.632~
$ black_pct  <dbl> 87.3239437, 14.2559834, 5.4430380, 3.7945396, 4.7696477, 2.8~
```

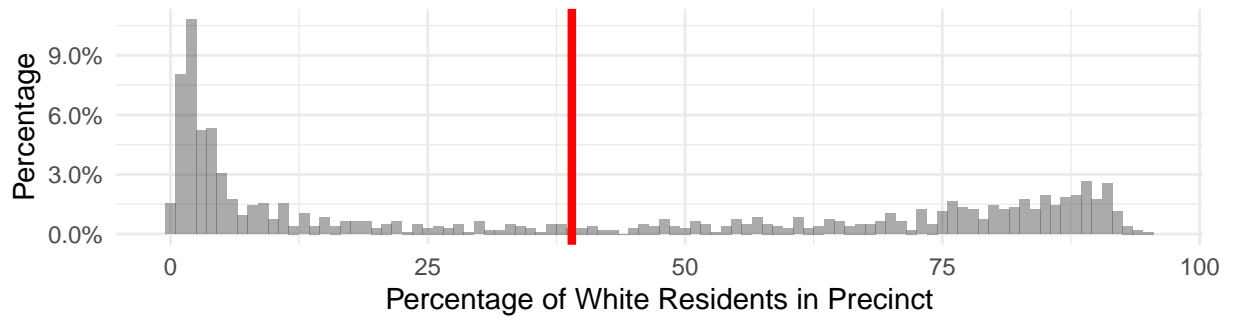
```
library(patchwork)

p1 <- wayne %>%
  ggplot() +
    geom_histogram(mapping = aes(white_pct, y = stat(density)),
                   binwidth = 1, alpha = 1/2) +
    theme_minimal() +
    scale_y_continuous(labels = percent_format()) +
    geom_vline(xintercept = 39, size = 1.5, colour = "red") +
    labs(x="Percentage of White Residents in Precinct",
         title="White Residents % by Precinct in Wayne County",
         y="Percentage", caption="Red vertical line is the mean.") +
    theme(
      plot.title = element_text(size = 12)
    )

p2 <- wayne %>%
  ggplot() +
    geom_histogram(mapping = aes(black_pct, y = stat(density)),
                   binwidth = 1, alpha = 1/2) +
    theme_minimal() +
    scale_y_continuous(labels = percent_format()) +
    geom_vline(xintercept = 50, size = 1.5, colour = "red") +
    labs(x="Percentage of Black Residents in Precinct",
         title="Black Residents % by Precinct in Wayne County",
         y="Percentage", caption="Red vertical line is the mean.") +
    theme(
      plot.title = element_text(size = 12)
    )

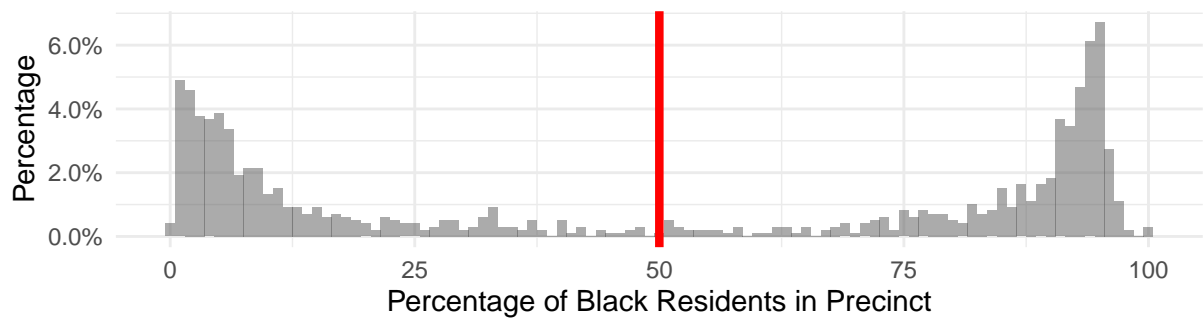
p1 / p2
```

White Residents % by Precinct in Wayne County



Red vertical line is the mean.

Black Residents % by Precinct in Wayne County



Red vertical line is the mean.

St. Louis City, MO Histogram

```
stl <- read_xlsx("stl_voting_precincts.xlsx")
glimpse(stl)
```

Rows: 223

Columns: 3

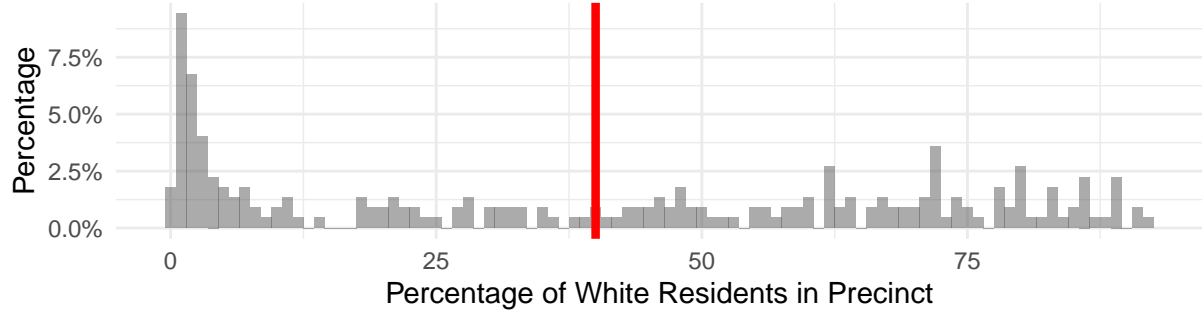
```
$ Label      <chr> "STL 1-2 Voting District, St. Louis city, Missouri", "STL 1--
$ white_pct  <dbl> 1.0777521, 3.4985423, 1.5497553, 2.3405973, 10.0783875, 0.62~
$ black_pct  <dbl> 95.53503, 92.41983, 97.30832, 92.09040, 85.94625, 93.16770, ~
```

```
p3 <- stl %>%
  ggplot() +
    geom_histogram(mapping = aes(white_pct, y = stat(density)),
                   binwidth = 1, alpha = 1/2) +
    theme_minimal() +
    scale_y_continuous(labels = percent_format()) +
    geom_vline(xintercept = 40, size = 1.5, colour = "red") +
    labs(x="Percentage of White Residents in Precinct",
         title="White Residents % by Precinct in St. Louis",
         y="Percentage", caption="Red vertical line is the mean.") +
    theme(
      plot.title = element_text(size = 12)
    )
```

```
p4 <- stl %>%
  ggplot() +
    geom_histogram(mapping = aes(black_pct, y = stat(density)),
                   binwidth = 1, alpha = 1/2) +
    theme_minimal() +
    scale_y_continuous(labels = percent_format()) +
    geom_vline(xintercept = 48, size = 1.5, colour = "red") +
    labs(x="Percentage of Black Residents in Precinct",
         title="Black Residents % by Precinct in St. Louis",
         y="Percentage", caption="Red vertical line is the mean.") +
    theme(
      plot.title = element_text(size = 12)
    )
```

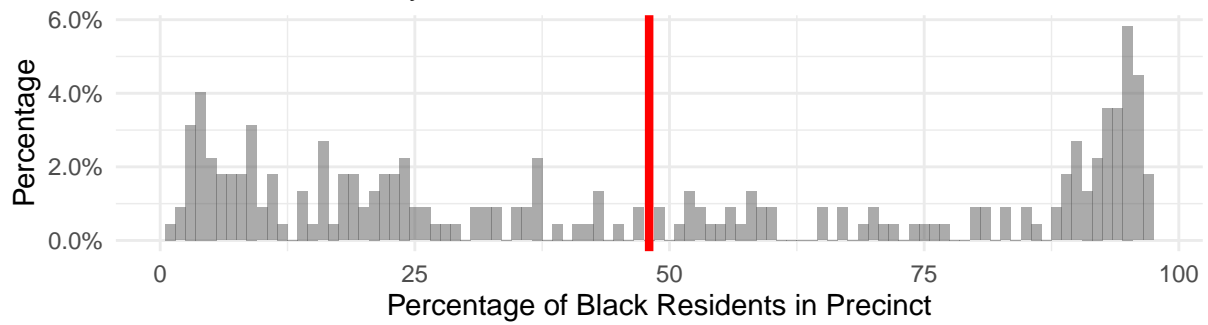
p3 / p4

White Residents % by Precinct in St. Louis



Red vertical line is the mean.

Black Residents % by Precinct in St. Louis



Red vertical line is the mean.